PLOS ONE

RESEARCH ARTICLE

Development and validation of smartwatch-based activity recognition models for rigging crew workers on cable logging operations

Eloise G. Zimbelman*☯, Robert F. Keefe☯

Department of Forest, Rangeland and Fire Sciences, University of Idaho, Moscow, ID, United States of America

☯ These authors contributed equally to this work.
* eloisez@uidaho.edu

Check for updates

OPEN ACCESS

Citation: Zimbelman EG, Keefe RF (2021) Development and validation of smartwatch-based activity recognition models for rigging crew workers on cable logging operations. PLoS ONE 16(5): e0250624. https://doi.org/10.1371/journal.pone.0250624

Editor: Chi-Hua Chen, Fuzhou University, CHINA

Received: December 11, 2020

Accepted: April 9, 2021

Published: May 12, 2021

Copyright: © 2021 Zimbelman, Keefe. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: A clear ethical concern restricts sharing our data set. Prior to implementation, the study was approved by the Institutional Review Board (IRB) at the University of Idaho and included developing an Informed Consent form to provide to all participants. We did not obtain permission from the participants to disclose any individual data during this written informed consent process. Although we recognize that the de-identification process suffices to practically meet the terms of anonymity for the purposes of publication in PLOS ONE, and for many other uses in medical research, public

Abstract

Analysis of high-resolution inertial sensor and global navigation satellite system (GNSS) data collected by mobile and wearable devices is a relatively new methodology in forestry and safety research that provides opportunities for modeling work activities in greater detail than traditional time study analysis. The objective of this study was to evaluate whether smartwatch-based activity recognition models could quantify the activities of rigging crew workers setting and disconnecting log chokers on cable logging operations. Four productive cycle elements (*travel to log*, *set choker*, *travel away*, *clear*) were timed for choker setters and four productive cycle elements (*travel to log*, *unhook*, *travel away*, *clear*) were timed for chasers working at five logging sites in North Idaho. Each worker wore a smartwatch that recorded accelerometer data at 25 Hz. Random forest machine learning was used to develop predictive models that classified the different cycle elements based on features extracted from the smartwatch acceleration data using 15 sliding window sizes (1 to 15 s) and five window overlap levels (0%, 25%, 50%, 75%, and 90%). Models were compared using multiclass area under the Receiver Operating Characteristic (ROC) curve, or AUC. The best choker setter model was created using a 3-s window with 90% overlap and had sensitivity values ranging from 76.95% to 83.59% and precision values ranging from 41.42% to 97.08%. The best chaser model was created using a 1-s window with 90% overlap and had sensitivity values ranging from 71.95% to 82.75% and precision values ranging from 14.74% to 99.16%. These results have demonstrated the feasibility of quantifying forestry work activities using smartwatch-based activity recognition models, a basic step needed to develop real-time safety notifications associated with high-risk job functions and to advance subsequent, comparative analysis of health and safety metrics across stand, site, and work conditions.

sharing of the data, even in anonymized form, specifically and clearly violates the terms under which participants willingly took part in the study. It is therefore unethical and inappropriate. These restrictions have been imposed by the University of Idaho IRB. Data requests may be sent to the University of Idaho IRB (contact via irb@uidaho.edu). All random forest models developed during the study have been made publicly available with metadata through the University of Idaho's Northwest Knowledge Network (NKN), which is an NSF-funded center for data management and storage. The models are available at: https://doi.org/10.7923/6g0x-ky24.

Funding: This work was funded by the University of Washington Pacific Northwest Agricultural Safety and Health (PNASH) Center (https://deohs.washington.edu/pnash/) Pilot Project subaward number UWSC10722 at the University of Idaho (RFK). The funders had no role in study design, data collection and analysis, or decision to publish. One project collaborator at PNASH provided minor feedback on a draft version of the manuscript. Comments did not involve changes to results or the decision to publish.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Cable logging operations consist of felling, yarding, processing, and loading work phases [1], with the yarding phase often characterized as six distinct cycle elements (outhaul, lateral out, hookup, lateral in, inhaul, and unhook) [2]. Many of the yarding tasks, such as pulling the cable laterally as well as hooking and unhooking logs, are manual, which can cause physiological strain [3]. Logging workers are often fatally injured when struck by objects such as falling trees, limbs, or machines [4–10]. Contact with objects and equipment accounted for 70.9% of fatalities among logging workers in 2017 [9] and 82.1% of fatalities among logging workers in 2018 [10]. Hand fallers and choker setters are particularly susceptible to these "struck-by" incidents, which accounted for 51.3% of injuries among hand fallers and choker setters between July 2010 and June 2015 in Montana and Idaho [8]. In an analysis of cable logging accidents, Tsioras et al. [11] found that broken spar and anchor trees, bouncing cables, and falling objects contributed to the majority of accidents and most incidents occurred when workers were struck by or struck against an object. While the increased mechanization of logging has generally led to a decrease in injury rates, ground crew members working alongside machines, rigging crew workers, and hand fallers are still at risk [4, 6, 8, 12] and may benefit from the use of a variety of positioning and wearable sensor-based technologies that increase situational awareness and reduce accidents.

Monitoring the current activities, safety status and location of individuals relative to workplace hazards on logging operations could be accomplished through (1) real-time location-sharing methods based on GNSS-RF (global navigation satellite system (GNSS) positioning paired with radio frequency (RF) transmission) for use in remote areas [13–17], (2) activity recognition modeling and incident detection, or (3) a combination of both. Human activity recognition involves using wearable sensors to distinguish between human physical activities. Most activity recognition models have been developed for everyday activities, such as walking, sitting, lying, standing, and other common physical movements referred to as activities of daily living (ADLs) [18–21], as well as for recreation and fitness applications [22–24]. Many smartphones and smartwatches are equipped with a variety of embedded sensors such as GNSS chips, accelerometers, gyroscopes, barometers, magnetometers, thermometers, decibel meters (microphones), and optical heart rate sensors [25–30]. Although a variety of purpose-built sensors have been developed, smartphones [26, 28, 31] and smartwatches [27, 29, 32–35] are popular for activity recognition modeling because they are ubiquitous and unobtrusive. Leveraging a variety of wearable and positioning sensors to develop occupational activity recognition models in forestry is a first step toward active monitoring that utilizes subsequent model predictions to help inform algorithms identifying falls or high-risk activities. Real-time prediction of work cycle elements represents an initial step toward informing smart, location- and activity-aware algorithms and alerts associated with detecting incidents and periods of elevated health and safety concern.

Development of activity recognition models generally consists of data collection, preprocessing, feature extraction and selection, and model development (Fig 1) [28, 36, 37]. Due to the advent of microelectromechanical systems (MEMS), inertial sensors have become smaller, more accurate, and less expensive and have been integrated into a variety of wearable sensors [38]. Data from these sensors is collected while users perform the activities of interest and is typically annotated with observed start and stop times. Preprocessing commonly consists of median filtering to remove noise spikes [39] and low pass or high pass filtering to isolate acceleration due to gravity from body acceleration [39–41]. To extract features for model development, a moving, or sliding, window is advanced through the dataset, defining subsets of the data from which relevant time (e.g., mean, median, variance, standard deviation, range,

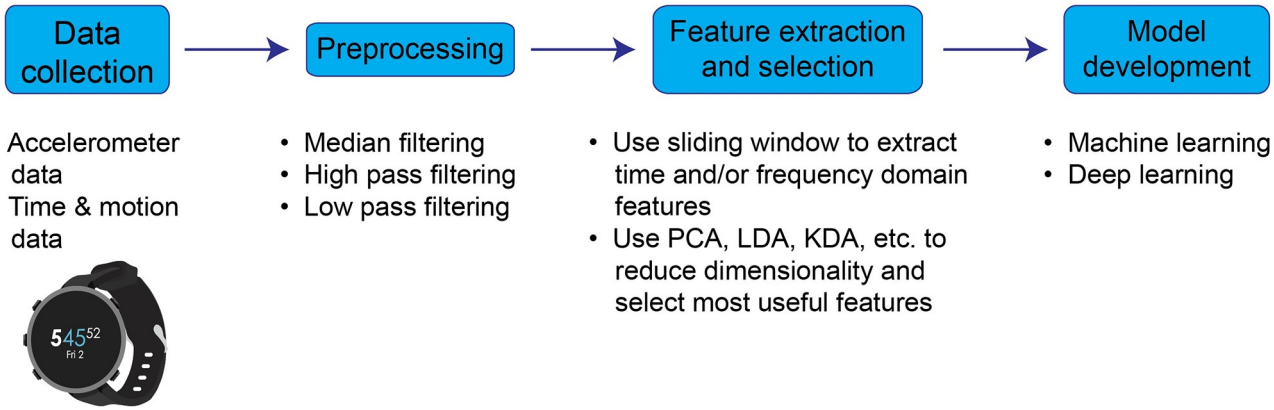

Fig 1. Outline of the general activity recognition model development process. Steps involved typically include (1) collecting time study data to pair with wearable sensor measurements, (2) preprocessing the data through filtering, (3) extracting time and/or frequency domain features using a sliding window and then selecting relevant features with which to build models, and (4) developing activity recognition models using machine learning or deep learning techniques. Ultimately, models may be programmed into apps on smartphones and smartwatches and subsequently used to characterize work activities in real-time to inform health and safety notifications.

https://doi.org/10.1371/journal.pone.0250624.g001

skewness and kurtosis) or frequency (e.g., Fast Fourier Transform and Discrete Transform coefficients) domain features are calculated [30, 31, 42, 43]. In order to reduce dimensionality and select the most useful features, a variety of techniques such as principal component analysis (PCA), singular value decomposition (SVD), linear discriminant analysis (LDA), or kernel discriminant analysis (KDA) can be used [41, 44, 45]. Finally, the extracted features are used to develop activity recognition models, often using machine learning algorithms such as Decision Trees, Random Forests (RFs), Support Vector Machines (SVMs), k-Nearest Neighbors (k-NN), Naïve Bayes, k-means, Hidden Markov Models (HMMs), Gaussian Mixture Models (GMMs), artificial neural networks (ANNs), and multilayer perceptron (MLP) [28, 31, 43, 46]. More recently, deep learning methods such as Restricted Boltzmann Machine, Autoencoders, Convolutional Neural Networks, and Recurrent Neural Networks have been shown to improve human activity recognition model performance compared to classical machine learning algorithms [47–50]. While deep learning can overcome some of the drawbacks of traditional machine learning by automatically extracting features and using more complex features [48–50], they are generally more computationally expensive and thus have not been widely implemented on resource-limited devices such as smartwatches and smartphones [47].

Model accuracy is thus affected by a variety of factors, such as the type and quality of sensors in the devices, sampling rate, device location on the body, machine learning algorithms employed for model development, features used for classification, and sliding window size used to extract these features [31, 42, 51]. In terms of window size, there is a tradeoff between detection times and recognition performance since smaller windows allow faster recognition speed [51] but longer windows have been shown to improve recognition performance for more complex, less repetitive activities [34]. In evaluating this tradeoff, Banos et al. [51] compared activity recognition models created with windows ranging in size from 0.25 s to 7 s and found that windows of 1–2 s provided the best accuracy while allowing for quick detection times. The amount of overlap between successive sliding windows also affects model accuracy, with larger overlap often leading to better model performance but at the cost of increased computational load [45]. Models with 50% overlap are common [20, 40, 52–55], but high accuracies can be obtained using nonoverlapping windows [29, 34, 51, 56]. Recently, "online" activity recognition, which refers to implementing the entire classification process (i.e., data

collection and pre-processing, feature extraction, and classification) locally on the device, has been investigated to attain near real-time classifications [31, 42]. Online recognition systems have been shown to have classification accuracies of 45.4–98.1% [42, 57, 58].

While most activity recognition models are specific to ADLs, the high-resolution data collected by these devices may also be useful for quantifying work activities. A limited number of studies have developed activity recognition for occupational activities based on wearable sensors, primarily for construction [59–63], while others have used visual observation or vision sensors to detect activities in the workplace [64, 65]. In natural resources, this approach has recently been proposed by Keefe et al. [30] and Pierzchała et al. [66] developed a method of automatically distinguishing between work phases in cable yarding by fusing data from multiple sensors. New GNSS-enabled smartwatches may offer lightweight alternatives to smartphone-based activity recognition models and may provide additional data that could supplement and improve these models. These types of wearable-based predictive models that quantify work activities on logging operations could inform loggers and equipment operators about their own or their coworkers' job activity status in near real-time, helping to increase situational awareness and safety on active timber sales. New, inexpensive mesh network communications technologies, such as those from goTenna Inc (Brooklyn, NY, USA), enable location- and data-sharing by connecting to smartphones via Bluetooth and allowing users to communicate through radio frequency. In addition to facilitating off-the-grid location-sharing, these devices may also be useful for sharing worker safety status derived from activity recognition models. The record of high-resolution data that results from activity recognition may also form the basis for quantifying occupational health and safety conditions in comparative analyses that span forest stand, site, and work conditions.

Time and motion studies [67] have been used extensively in conventional forest operations research to quantify productive work cycle elements and delay [2, 68–73]. By defining and analyzing the individual work cycle elements performed by equipment or by individuals engaged in motor-manual operations, an objective of time study research is often to identify opportunities for improving occupational production rates and reducing delay time [1, 2, 71, 72, 74]. Time study analyses are used as the basis for regression [1, 2, 69, 71, 74, 75] and machine learning [76, 77] models that predict work cycle time as a function of stand or site conditions. In order to estimate logging costs per unit wood volume, machine rate estimates determined using methods outlined by Miyata [78] and Brinker et al. [79] are paired with these cycle time prediction models [1, 71, 74, 75, 80]. In recent years, GNSS has been used in time studies to automate the estimation of cycle times [81], calculate machine productivity [82], characterize machine movements [83], and improve operational monitoring [84]. Additionally, time and motion studies have been conducted using both GNSS receivers and accelerometers to monitor tree planting [85], characterize manual felling using brush cutters [86], distinguish between chipping tasks [87], and monitor tilt and motion of various harvesters and forwarders in order to analyze operating conditions [88]. The availability of high-resolution sensor data collected seamlessly from GNSS-enabled mobile and wearable devices in real-time provides an opportunity to further model forestry work activities in greater detail than has been done using traditional methods, while simultaneously providing the basis for improved characterization of digital health and safety.

Prior research evaluating use of wearable sensors to monitor and model forestry work activities includes a small body of recent literature. Fitness and sleep bands have been used (1) to monitor the physical activity and sleep patterns of forestry workers in order to understand how these factors may contribute to workplace hazards [89], and (2) to predict forestry worker fatigue by comparing heart rate and step count data to reaction and decision-making times [90]. Smartwatches paired with heart rate monitor chest straps have been used to evaluate

workload associated with manual tree felling [91], while external accelerometers attached to machines have been used to develop ANNs that classify the activities of manually-driven bandsaws [92] and recognize activities associated with manual felling [93]. Preliminary activity recognition models have been developed for cable yarding work phases using a combination of smartphone sensor (global positioning system (GPS) and inertial measurement unit (IMU)) and camera data [66]. However, activity recognition models have yet to be developed for other forestry work positions, such as rigging crew workers setting and disconnecting log chokers. Furthermore, smartwatches have not previously been used in forestry activity recognition, so it is unknown how the prediction accuracy of models developed using these devices will compare to smartphone-based models. Specifically, smartwatches may record different movement patterns than smartphones due in part to different device locations on the body (i.e., wrist vs. hip) [27, 34]. In this study, we developed smartwatch-based activity recognition models for rigging crew workers on cable logging operations in order to address two specific research objectives. Our first objective was to develop models that predict choker setter rigging crew work activities with at least 80% sensitivity. Our second objective was to develop models that predict chaser rigging crew work activities with at least 80% sensitivity. Random forest machine learning was used to develop watch-based activity recognition models based on time and motion study data collected on five active timber sales in North Idaho, USA. Model accuracy was calculated based on the percent of time work elements were predicted correctly.

Materials and methods

Ethics statement

Fourteen loggers voluntarily participated in this study. Prior to data collection, the experimental protocol was approved by the University of Idaho Institutional Review Board (IRB number: 18–202). Participants received both oral and written information regarding the study design and provided their written, informed consent. Participants were selected based on available, operational cable-logging activities occurring in the North Idaho region during the sampling period and reflect the general demographics of the study population. While we did not collect demographic information from our participants, recent sampling in the region has shown that 55.4% of the logging workforce in Idaho is 50 or more years of age [17]. However, rigging crew workers are generally younger than the median age.

Data collection and processing

Time and motion study (i.e., observational elemental time analysis) data was collected in conjunction with GNSS watch sensor data using two days of sampling on each of five timber sales. Timber sales occurred on state and industrial cable logging operations. Choker setters (who are responsible for setting chokers on logs to be yarded) and chasers (who are responsible for disconnecting chokers from yarded logs) were observed visually. Four productive cycle elements (*travel to log*, *set choker*, *travel away*, *clear*) were timed for the choker setters and four productive cycle elements (*travel to log*, *unhook*, *travel away*, *clear*) were timed for the chasers (Table 1). For the choker setter activities, *travel to log* began when the choker setter started walking toward the carriage to grab the chokers. *Set choker* began when the choker setter arrived at a log and began preparing chokers. *Travel away* began when the choker setter finished setting chokers and started to walk away from the log. *Clear* began when the choker setter stopped walking away from the log and was "in the clear". For the chaser activities, *travel to log* began when the chaser started walking toward the landed logs. *Unhook* began when the chaser reached for the chokers to begin unhooking them. *Travel away* began when the chaser finished unhooking the chokers and started to walk away from the log(s). *Clear* began when

Table 1. Summary of productive cycle elements for choker setter and chaser work activities.

Position	Activity	Activity begins when subject:
Choker setter	Travel to log	Initiates walking toward carriage to acquire chokers
	Set choker	Arrives at log
	Travel away	Finishes setting choker
	Clear	Stops walking away when safely "in the clear"
Chaser	Travel to log	Initiates walking toward the landed logs
	Unhook	Reaches for the chokers to begin unhooking
	Travel away	Finishes unhooking chokers
	Clear	Stops walking away when safely "in the clear"

https://doi.org/10.1371/journal.pone.0250624.t001

the chaser stopped walking away from the logs and was "in the clear". For both the choker setter and chaser, *clear* included everything workers did outside of the other three work elements. In order to develop a generalized model for functional use, minor delay events were included within the relevant work element during which they occurred. The clock on a Google Pixel smartphone was used to record the true start and stop times for each work activity cycle using the TimeStamp application (version 0.4.0) [94]. Workers wore Garmin Fenix 5S Plus watches, which record GNSS locations, heart rate, and raw accelerometer data, on their non-dominant wrist. All sensor data was recorded on the Garmin watches using the RawLogger application (version 1.0.20190520a) [95] from Garmin Connect. The accelerometer sensor data was collected at a 25-Hz frequency, while other sensors were recorded at a 1-Hz frequency, which are the default frequencies within the RawLogger application. All sensor data was exported as a Garmin FIT file and subsequently converted into a *.csv* file. Only the watch accelerometer data, collected at 25 Hz and recorded in thousandths of a gravity (mgn), was used in model development.

All data processing, analysis and model development was done in the R statistical programming environment, version 4.0.0 [96]. After data collection, all observations in the datasets were labeled according to the manually recorded start and stop times. Specifically, each observation whose timestamp fell within the start and stop time for a particular activity cycle was assigned a label for that activity element (i.e., *travel to log*, *set choker*, *travel away*, *clear*, etc.). Delay times were included as the corresponding productive work element because the majority of delays fell within *clear*, included a diverse range of physical movements associated with the workers, and because of the intended final use of a general model in continuous, real-time prediction. After labeling, the raw acceleration values (in the x, y, and z dimensions) were filtered using a Finite Impulse Response (FIR) bandpass filter of order 8. Filter band edges were 0.5 and 0.9. Rather than using the x, y, and z values of the acceleration sensor, the acceleration magnitude was calculated using Eq (1) and used in an effort to reduce the effects of orientation on recognition performance:

$$A_{mag} = \sqrt{A_x^2 + A_y^2 + A_z^2} \tag{1}$$

Where A_{mag} is the filtered overall acceleration magnitude, and A_x, A_y, and A_z are the filtered acceleration sensor values in the x, y, and z dimensions, respectively.

Activity recognition model development

Ten time domain features (mean, standard deviation, maximum, minimum, median absolute deviation, mean absolute deviation, skewness, interquartile range, range, and kurtosis) were extracted from the filtered acceleration magnitude values from both the choker setter and

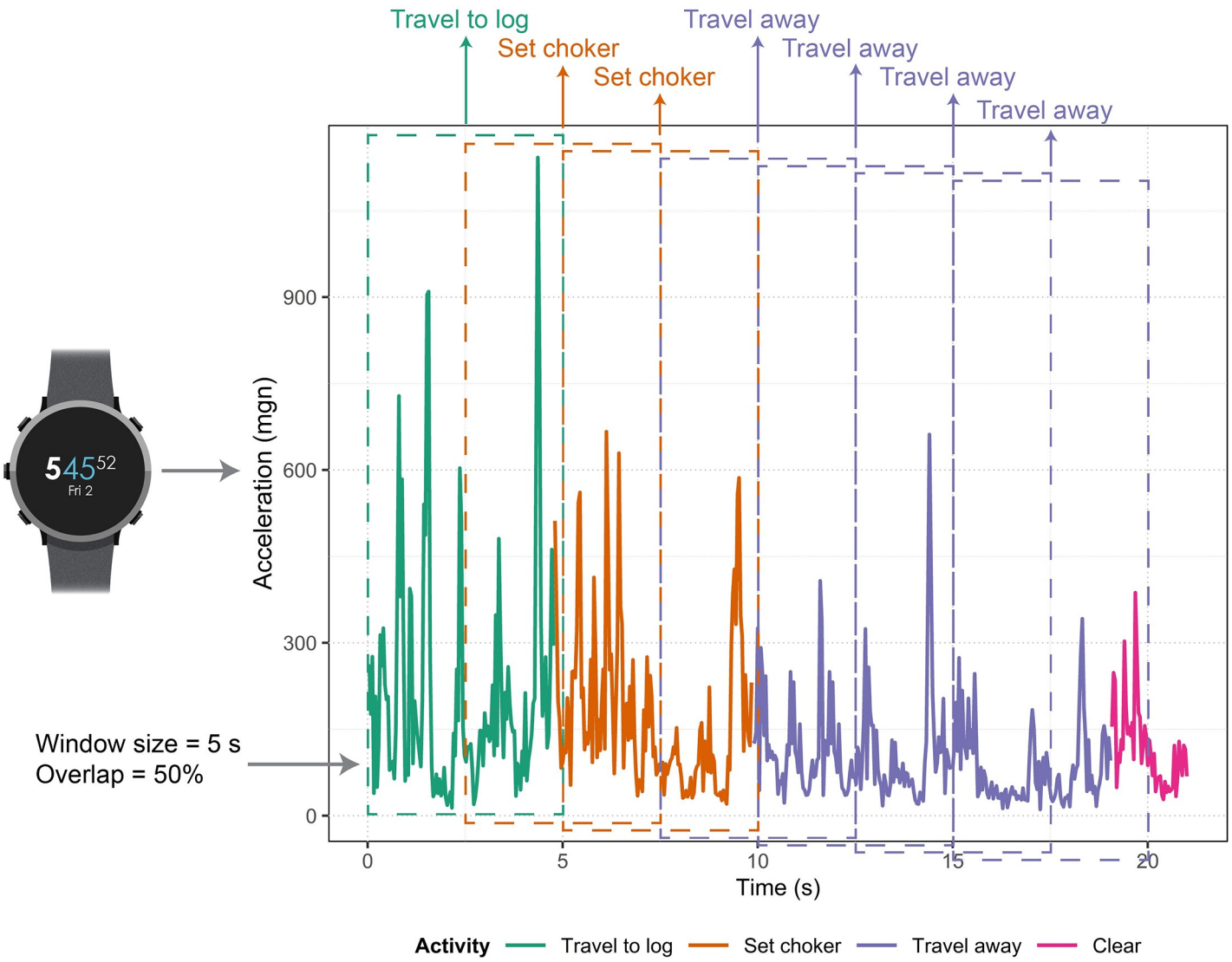

Fig 2. Overview of a hypothetical choker setter activity recognition model running on a smartwatch. The activity recognition model depicted is using a 5-s window with 50% overlap to predict the four work activities. The figure shows filtered acceleration magnitude data, which is colored according to the actual work cycles. Each time a window (shown as rectangles with dashed lines) is used to extract features, the model predicts the work cycle (shown as labels above the windows).

https://doi.org/10.1371/journal.pone.0250624.g002

chaser work activity data using 15 different sizes of sliding windows (ranging from 1 to 15 s). For example, using a 3-s window and data recorded at 25 Hz, features were calculated using the previous 75 observations (representing 3 s of data) each time the window was advanced. Windows with 0%, 25%, 50%, 75%, and 90% overlap were used, resulting in five feature extraction methods for each window size. For instance, using 25% overlap meant that the next window did not begin until the current window was 75% complete. After filtering and applying sliding windows, the resulting datasets were separated into 2/3 training and 1/3 testing data. Data was separated randomly, but the relative ratios of each activity were preserved because the data was highly imbalanced. The randomForest function in the R randomForest package (version 4.6–14) [97] was used to create random forest models to predict the four work cycle elements of both the choker setter and chaser based on the sensor measurements (Fig 2). Because the data was imbalanced, models were created using stratified sampling according to

activity, with sample size based on the number of instances of the least common activity. In terms of the choker setter models, the least common activity was *travel away*. The least common chaser activity was *travel to log*. Random forest models can be tuned via a variety of parameters, such as the number of trees to grow (n_{tree}) and the number of predictor variables randomly selected at each node (m_{try}) [97, 98]. In this study, random forest models were created using 150 trees, since previous work has suggested that using 64–128 trees is appropriate for balancing performance and processing time [99]. Because varying m_{try} generally does not have a significant effect on model performance [97, 100], models in this study were built with the default value of m_{try} (the square root of the total number of variables). The relationship between the number of trees and model accuracy was evaluated using the out-of-bag (OOB) sample error rates calculated internally by the random forest algorithm.

Initially, choker setter and chaser models were created for the 15 window sizes using 90% overlap and all ten features, but model accuracy was poor. Thus, principal component analysis (PCA) was used to reduce the number of features in the models. The same ten time domain features described above were calculated for each activity cycle for both the choker setter and chaser datasets using the filtered acceleration magnitude values. PCA was performed on these ten features for both the choker setter and chaser datasets. Principal components (PCs) that accounted for $> 95\%$ of the variation in each dataset were considered and individual variables with loadings $> |0.4|$ within these PCs were used as predictors in the final models.

After selecting model predictors, a total of 75 choker setter models and 75 chaser models were created using the 15 window sizes for each of the five levels of overlap. Models were created using the training datasets and the confusionMatrix function in the R caret package (version 6.0–85) [101] was used to calculate a variety of model accuracy metrics based on the testing datasets. All models were initially evaluated using sensitivity, specificity, and precision to compare the effects of overlap levels and window sizes. Sensitivity, specificity, and precision were calculated using Eqs (2–4):

$$Se = \frac{TP}{TP + FN} \tag{2}$$

$$Sp = \frac{TN}{TN + FP} \tag{3}$$

$$Pr = \frac{TP}{TP + FP} \tag{4}$$

Where *Se* is the sensitivity, *Sp* is the specificity, and *Pr* is the precision. *TP* is the number of true positives (i.e., the number of correctly classified instances of a given class), *TN* is the number of true negatives (i.e., for a given class, the number of instances of all other classes that are classified as anything other than the class of interest), *FP* is the number of false positives (i.e., the number of instances that are incorrectly classified as belonging to a given class), and *FN* is the number of false negatives (i.e., the number of instances of a given class that are incorrectly classified as a different class). Sensitivity, specificity, and precision were then converted to and reported as percentages. Sensitivity, or recall, is the true positive rate and represents the percentage of correctly identified activities of a particular class [102]. Specificity is the true negative rate and measures the percentage of correctly detected negative occurrences of a particular class [102]. Precision, or positive predictive value, measures the percentage of detected instances of an activity that represents a real occurrence [102]. Finally, the multiclass area under the Receiver Operating Characteristic (ROC) curve, or AUC, was calculated for each model. The AUC corresponds to the probability that a classifier will rank a randomly chosen

positive instance higher than a randomly chosen negative instance, with higher AUC values indicating better performance [103]. AUC is a common criterion for evaluating the performance of classification algorithms [104, 105]. It has also been shown to be relatively robust to data imbalance [106–108]. The multiclass.roc function in the R pROC package (version 1.16.2) [109] was used to compute the multiclass AUC according to the method defined by Hand and Till [110]. The multiclass AUC value was then used to compare models and choose the best window size and overlap level for the choker setter and chaser models. Final models were evaluated based on the three metrics described previously (sensitivity, specificity, and precision) as well as F_1 values and balanced accuracy. F_1 values and balanced accuracy were calculated using Eqs (5 and 6):

$$F_1 = \frac{2 * Pr * Se}{Pr + Se} \tag{5}$$

$$BA = \frac{Se + Sp}{2} \tag{6}$$

Where F_1 is the F_1 value and BA is the balanced accuracy. Pr is the precision, Se is the sensitivity, and Sp is the specificity. The F_1 value represents the harmonic mean of precision and recall (sensitivity) and is generally thought to be more robust when dealing with imbalanced classes [51]. It ranges from zero to one, with zero representing no capacity for recognition and one corresponding to perfect recognition [51]. The F_1 value was calculated using the rate of precision and sensitivity (rather than the percent). Balanced accuracy is simply the mean of sensitivity and specificity and was calculated using the percentage values of these two metrics.

Results

Work activity cycle times

For the choker setter work activities, *travel to log* averaged 17.27 s (s = 12.30 s), *set choker* averaged 19.80 s (s = 16.06 s), *travel away* averaged 13.11 s (s = 6.83 s), and *clear* averaged 220.31 s (s = 519.17 s) (Table 2). For the chaser work activities, *travel to log* averaged 4.59 s (s = 4.42 s), *unhook* averaged 8.00 s (s = 6.42 s), *travel away* averaged 6.69 s (s = 3.73 s), and *clear* averaged 264.48 s (s = 481.49 s) (Table 2). Mean work activity cycle times for the chaser were generally shorter than the mean cycle times for the choker setter, and the maximum cycle times tended to be slightly longer for the choker setter activities compared to the chaser activities (Table 2). When expressed as a percentage of the mean elemental time for individual work cycles other than *clear*, the coefficient of variation (CV) ranged from 52.11% to 81.09% for the choker setter and from 55.70% to 96.28% for the chaser (Table 2). *Clear* was the only element for which the CV > 100% for both the choker setter and chaser (Table 2).

Table 2. Summary statistics (in seconds) of cycle times for choker setter and chaser work activities.

Position	Activity	Mean (s)	SD (s)	CV (%)	Range (s)	Median (s)	1st Quartile (s)	3rd Quartile (s)
Choker setter	Travel to log	17.27	12.30	71.23	0.72–105.56	13.96	9.06	21.82
	Set choker	19.80	16.06	81.09	1.14–188.21	14.73	9.46	24.61
	Travel away	13.11	6.83	52.11	2.5–55.74	11.87	8.74	15.44
	Clear	220.31	519.17	235.65	5.7–8441.34	120.22	93.55	166.14
Chaser	Travel to log	4.59	4.42	96.28	0.54–89.92	3.85	2.74	5.49
	Unhook	8.00	6.42	80.31	0.86–71.58	6.52	4.17	10.07
	Travel away	6.69	3.73	55.70	1.19–26.14	5.61	3.95	8.85
	Clear	264.48	481.49	182.06	4.93–8238.2	174.55	144.49	229.41

https://doi.org/10.1371/journal.pone.0250624.t002

Participant data

Due to the observational nature of data collection, the amount of data used to train and test the random forest models varied between participants. The average amount of data per choker setter ranged from 0.80 hrs (s = 0.01 hrs) to 6.74 hrs (s = 0.05 hrs) for training and from 0.40 hrs (s = 0.01 hrs) to 3.31 hrs (s = 0.05 hrs) for testing. The average amount of data per chaser ranged from 0.80 hrs (s = 0.02 hrs) to 4.59 hrs (s = 0.04 hrs) for training and from 0.39 hrs (s = 0.02 hrs) to 2.26 hrs (s = 0.04 hrs) for testing. When expressed as a percentage of the mean training and testing sample times for individual choker setters, the standard deviation ranged from 0.69% to 1.65% of the training sample times and from 1.39% to 3.34% of the testing sample times. Similarly, for individual chasers, the standard deviation ranged from 0.93% to 2.04% of the training sample times and from 1.90% to 4.10% of the testing sample times.

Principal component analysis

In terms of the choker setter, the first PC accounted for 98.13% of the variation in the dataset and the only variables with loadings > |0.4| in the first PC were the acceleration maximum and range (Table 3). Similarly, in terms of the chaser, the first PC accounted for 97.92% of the variation in the dataset and the only variables with loadings > |0.4| in the first PC were the acceleration maximum and range (Table 3). Thus, acceleration maximum and range were selected as predictors in the final models for both workers. Biplots of the first two PCs for both the choker setter and chaser datasets illustrate the strong effect of acceleration maximum and range on the two datasets (Fig 3).

Number of trees

At the 90% overlap level, the classification accuracies determined internally by the random forest algorithm for both the choker setter and chaser models leveled off after 25–50 trees for most window sizes (Fig 4), and similar trends were observed for the other overlap levels (not shown). This suggests that building our models with 150 trees was sufficient.

Table 3. Summary of choker setter and chaser PCA results.

	Position	**Choker setter**				**Chaser**			
	PC	**PC1**	**PC2**	**PC3**	**PC4**	**PC1**	**PC2**	**PC3**	**PC4**
	Percent of Variance	***98.13%***	***1.77%***	***0.05%***	***0.03%***	***97.92%***	***1.83%***	***0.14%***	***0.08%***
Variable	Mean	-0.038	**0.554**	0.235	-0.303	-0.030	**0.527**	-0.221	0.029
	Standard deviation	-0.054	0.398	**-0.686**	-0.353	-0.053	**0.443**	0.395	**-0.624**
	Maximum	**-0.705**	-0.036	0.056	-0.062	**-0.705**	-0.024	-0.014	0.028
	Minimum	0.000	0.039	0.092	-0.158	0.000	0.041	-0.042	-0.002
	Median absolute deviation	-0.021	0.356	0.354	0.163	-0.014	0.321	-0.299	0.285
	Mean absolute deviation	-0.031	0.329	-0.264	-0.063	-0.028	0.349	0.125	-0.278
	Skewness	-0.001	-0.009	0.010	-0.036	-0.001	-0.011	-0.020	-0.022
	Interquartile range	-0.035	**0.524**	0.296	0.345	-0.025	**0.501**	-0.323	0.284
	Kurtosis	-0.009	-0.131	**0.422**	**-0.771**	-0.020	-0.208	**-0.763**	**-0.607**
	Range	**-0.705**	-0.074	-0.036	0.096	**-0.705**	-0.064	0.028	0.030

The percentages of variance for the first four PCs as well as the individual loadings for each variable for the first four PCs are shown. Numbers in **bold** indicate variables with loadings > |0.4|.

https://doi.org/10.1371/journal.pone.0250624.t003

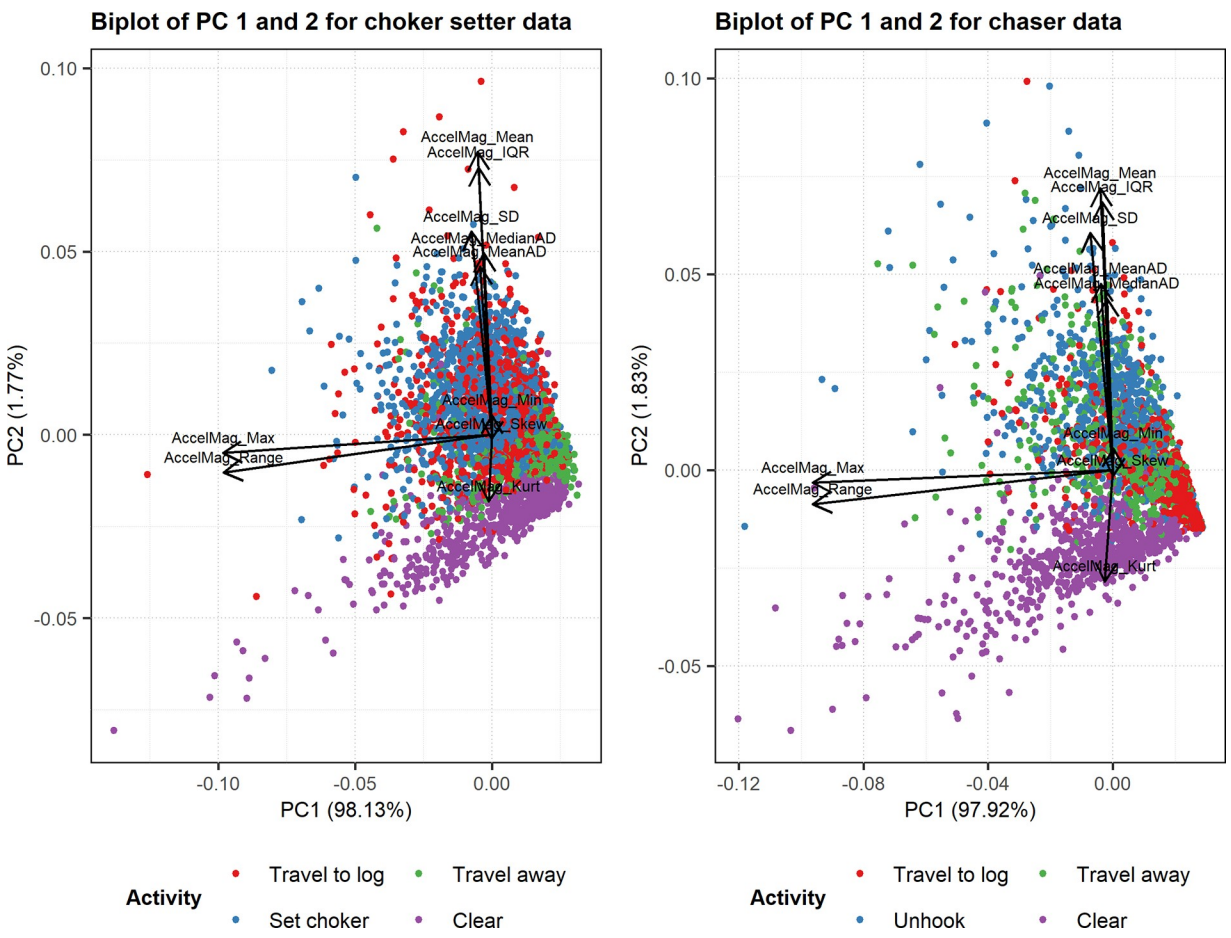

Fig 3. Biplots of PCs 1 and 2 for the choker setter and chaser datasets. The color of points on each plot indicates work cycle element categories.

https://doi.org/10.1371/journal.pone.0250624.g003

Choker setter models

The sensitivity, specificity, and precision of the choker setter models were highest when using sliding windows with 90% overlap for all activities and window sizes (Fig 5). Using sliding windows with 75% overlap resulted in the second-highest values of these same metrics for all activities and window sizes (Fig 5). The remaining overlap levels (50%, 25%, and 0%) resulted in the lowest values of these metrics (Fig 5). Finally, sensitivity, specificity, and precision generally did not vary noticeably between window sizes for most activities (Fig 5).

Chaser models

The sensitivity, specificity, and precision of the chaser models were highest when using sliding windows with 90% overlap for all activities and window sizes (Fig 6). Using sliding windows with 75% overlap resulted in the second-highest values of these same metrics for many activities and window sizes (Fig 6). The remaining overlap levels (50%, 25%, and 0%) generally resulted in the lowest values of these metrics (Fig 6). Additionally, sensitivity tended to decrease as window size increased for most activities (Fig 6, first row) while specificity decreased slightly with increasing window size for the *travel to log* activity but did not show much variation with window size for the other three activities (Fig 6, second row). Precision

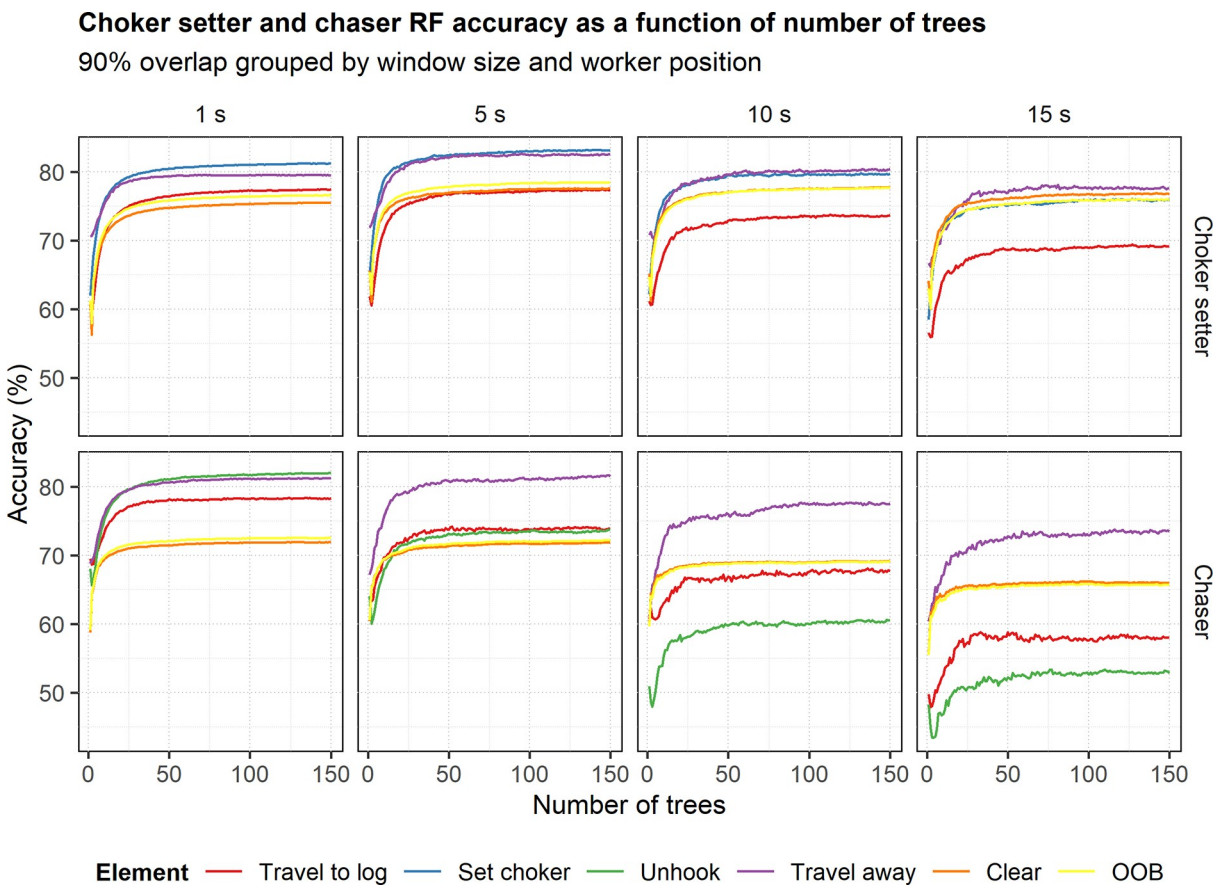

Fig 4. Choker setter and chaser random forest model accuracy as a function of the number of trees. The plots are grouped by worker type (choker setter or chaser) and window size. Line color indicates overall model (OOB) accuracy as well as accuracy for the work cycle elements. Only the 90% overlap of the 1-, 5-, 10- and 15-s windows are shown.

https://doi.org/10.1371/journal.pone.0250624.g004

decreased with increasing window size for the *travel to log* and *unhook* activities at the higher overlap levels but did not vary as noticeably for the *travel away* and *clear* activities (Fig 6, third row).

Model selection

The multiclass AUC plot indicates that the 90% overlap level resulted in the highest AUC values across all window sizes for both the choker setter and chaser models (Fig 7). AUC did not vary much with window size for the choker setter models (Fig 7, top row). For the chaser models, AUC decreased with increasing window size for the 90% and 75% overlap levels but varied less with window size for the other overlap levels (Fig 7, bottom row). A 3-s window with 90% overlap had the highest AUC (94.42%) for the choker setter models (Fig 7, top row) and a 1-s window with 90% overlap had the highest AUC (93.62%) for the chaser models (Fig 7, bottom row). Thus, these two models were chosen as the optimal choker setter and chaser models.

Sensitivity for the selected choker setter model (3-s window, 90% overlap) ranged from 76.95% to 83.59% for the four activities (Table 4). Precision for this model ranged from 41.42% to 97.08% for the four activities, specificity ranged from 91.36% to 94.73%, F_1 values ranged from 0.55 to 0.86, and balanced accuracy ranged from 85.02% to 88.90% (Table 4). The confusion matrix for the selected choker setter model (3-s window, 90% overlap) illustrates

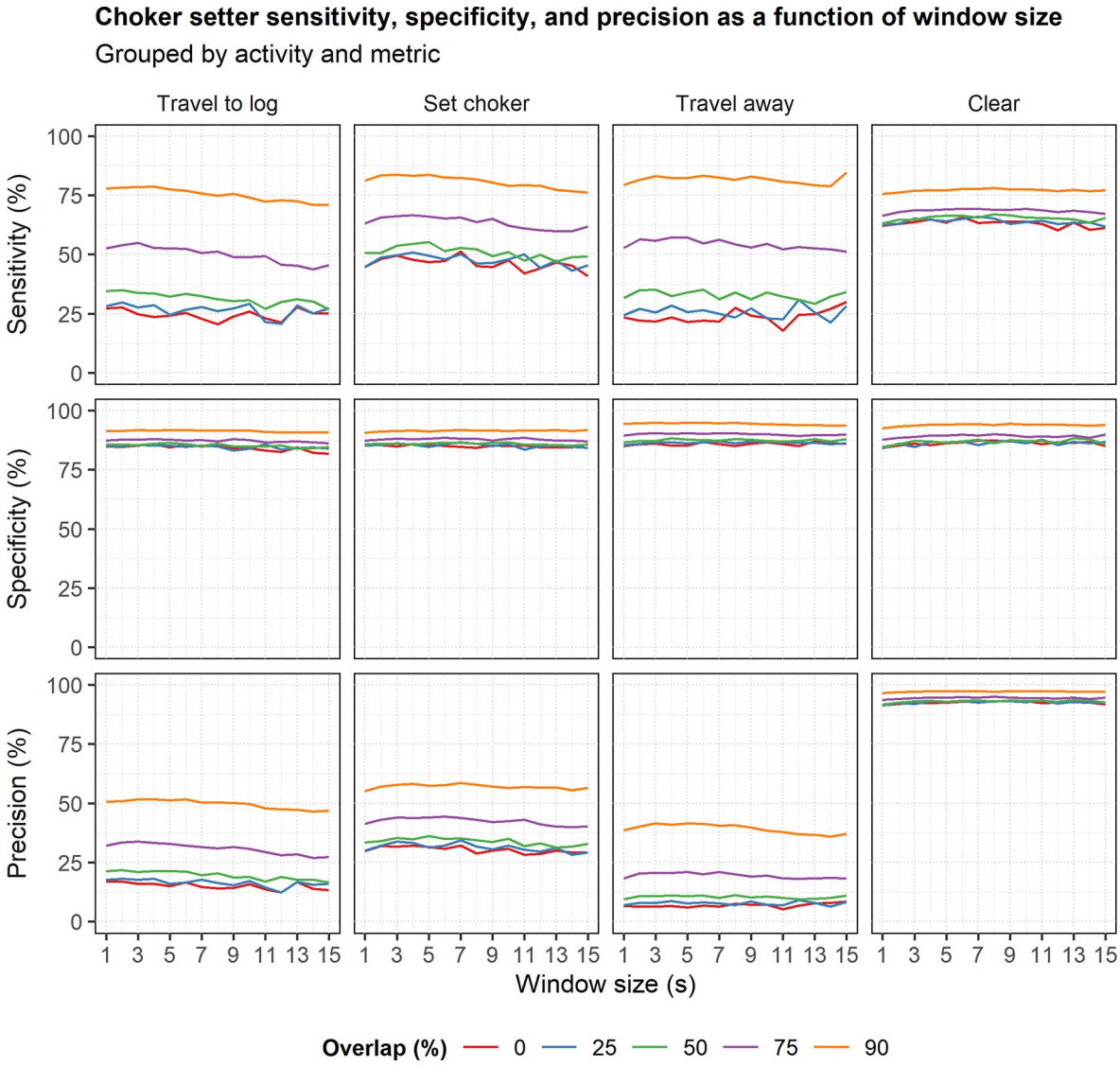

Fig 5. Choker setter sensitivity, specificity, and precision as a function of window size. The plots are grouped by metric and work activity. Line color indicates window overlap level.

https://doi.org/10.1371/journal.pone.0250624.g005

that the *set choker* activity was most often confused with *travel to log* (Table 5). *Travel away* was somewhat equally confused with the other three activities (*set choker*, *travel to log*, and *clear*), while *travel to log* was most often confused with either *set choker* or *clear* (Table 5). Finally, *clear* was most often confused with either *set choker* or *travel to log* (Table 5).

Sensitivity for the selected chaser model (1-s window, 90% overlap) ranged from 71.95% to 82.75% for the four activities (Table 6). Precision for this model ranged from 14.74% to 99.16% for the four activities, specificity ranged from 89.76% to 92.43%, F_1 values ranged from 0.25 to 0.83, and balanced accuracy ranged from 81.97% to 86.54% (Table 6). The confusion matrix for the selected chaser model (1-s window, 90% overlap) shows that *clear* was most often confused with *travel away* but was also frequently mistaken for *unhook* (Table 7). *Travel away* was most often confused with either *unhook* or *clear*, while *travel to log* was most often mistaken for *clear* (Table 7). Finally, *unhook* was most often confused with either *clear* or *travel away* (Table 7).

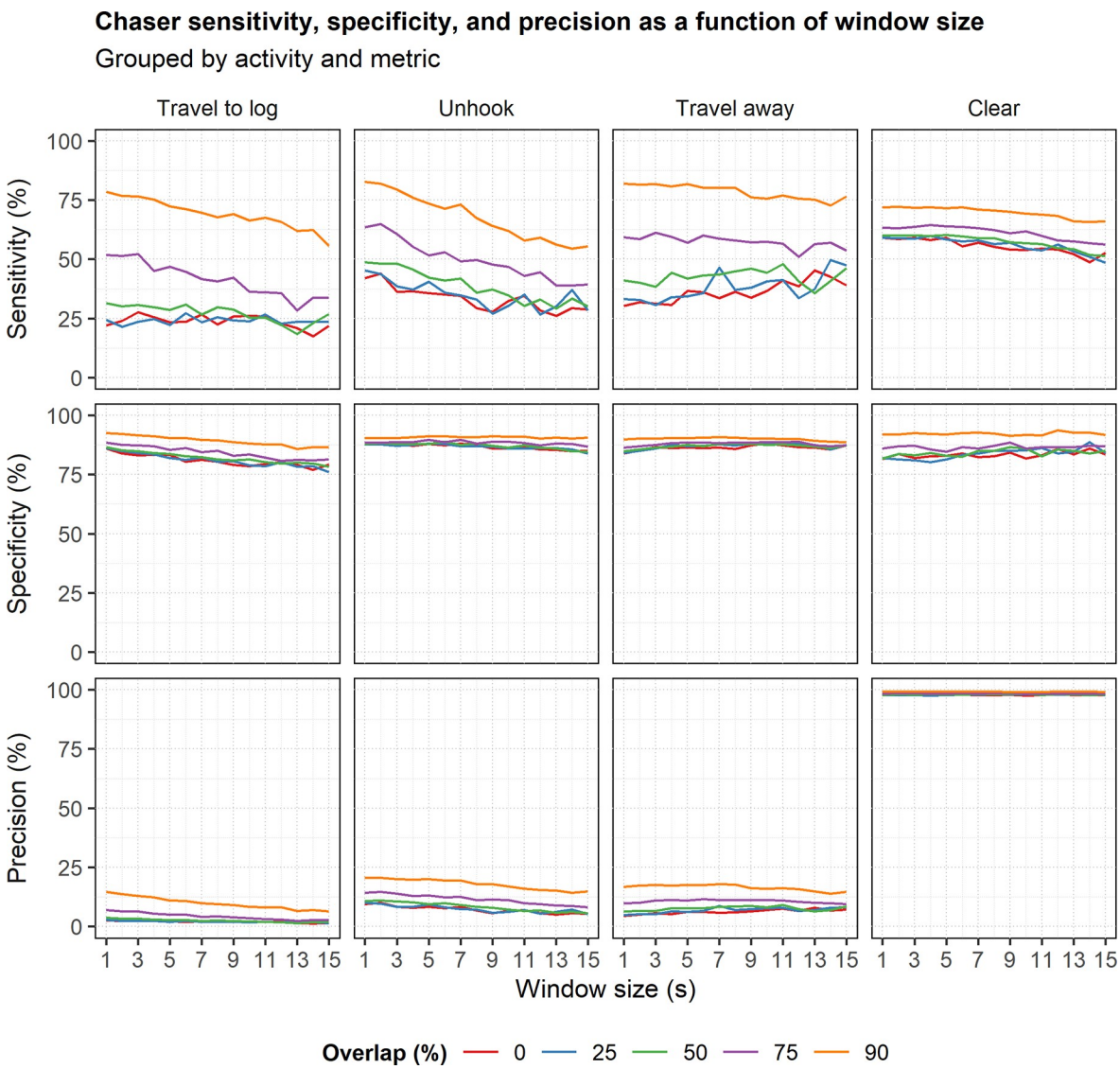

Fig 6. Chaser sensitivity, specificity, and precision as a function of window size. The plots are grouped by metric and work activity. Line color indicates window overlap level.

https://doi.org/10.1371/journal.pone.0250624.g006

Discussion

Our results show that activity recognition models based on smartwatch accelerometers can characterize work activities for rigging crew workers setting and disconnecting log chokers on cable logging operations, with the best model sensitivities ranging from 76.95% to 83.59% for choker setters and from 71.95% to 82.75% for chasers. While not all activities met our objective of 80% sensitivity, these values are consistent with models based on smartwatches and wrist-worn accelerometers developed for other activities in previous studies [27, 29, 32, 33, 35, 54–56]. The benchmark of 80% sensitivity was established ahead of time as part of project development. While we recognize that a combination of metrics may be more suitable for future use, particularly when dealing with imbalanced datasets, this was a pilot study intended to help establish methods prior to a larger modeling effort.

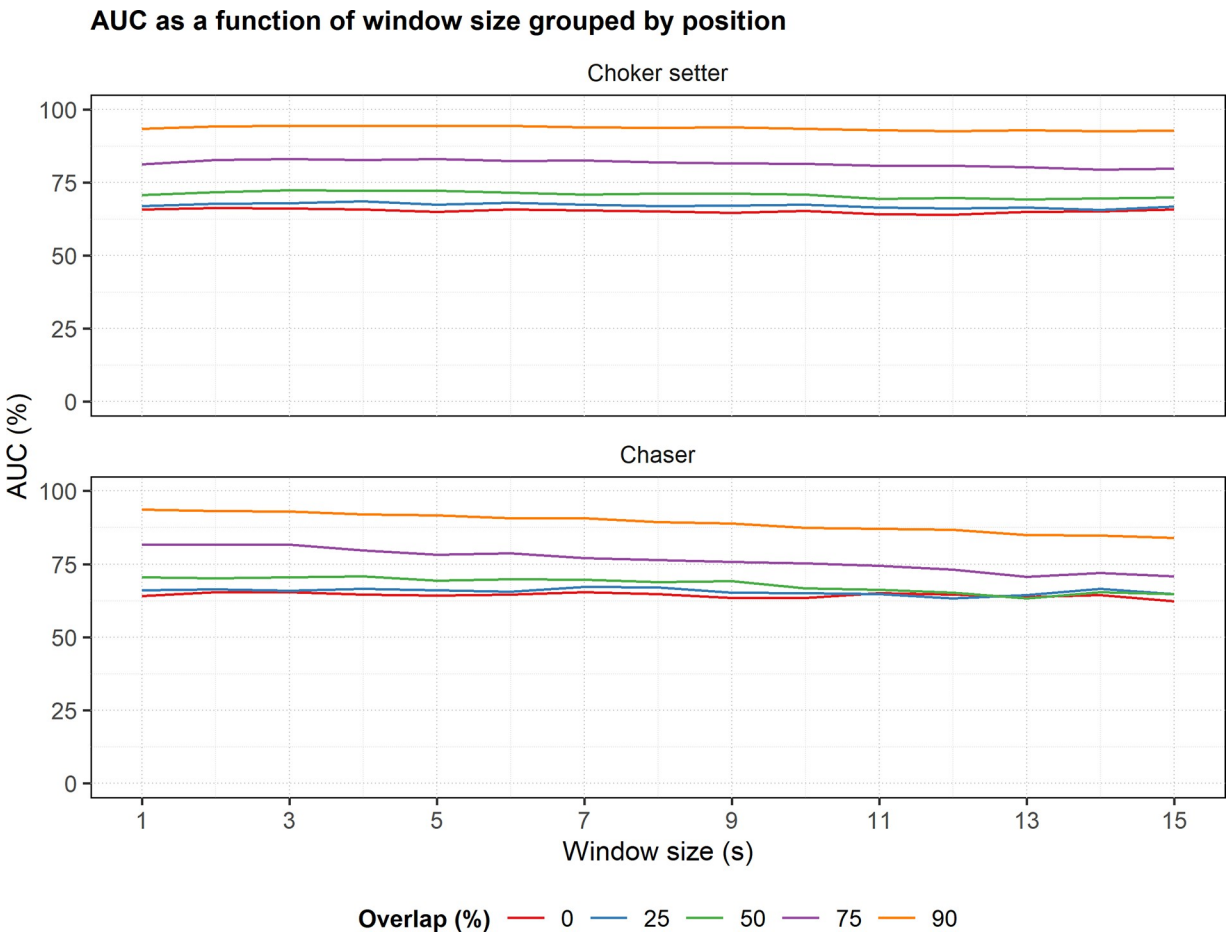

Fig 7. Choker setter and chaser multiclass AUC as a function of window size. The plots are grouped by worker type (choker setter or chaser). Line color indicates window overlap level.

https://doi.org/10.1371/journal.pone.0250624.g007

In contrast to the fairly high sensitivity values, precision was poor for most activities for both the choker setter and chaser models. This may be due to the imbalanced nature of the data. The proportions of choker setter activities were 10.26% *travel to log*, 12.41% *set choker*, 4.29% *travel away*, and 73.04% *clear*. The proportions of chaser activities were 1.64% *travel to log*, 2.95% *unhook*, 2.45% *travel away*, and 92.95% *clear*. Furthermore, the high CV values in Table 2 suggest there was high variability within individual work cycle elements, especially within the *clear* activity for both the choker setters (CV = 235.65%) and chasers (CV = 182.06%), which likely resulted at least in part from the inclusion of delay time in work elements. Additional variability may have been introduced to the *clear* activity due to the fact that *clear* encompassed everything workers did outside of the other three work elements. The

Table 4. Accuracy metrics for the best choker setter model (created with a 3-s window and 90% overlap).

Activity	Sensitivity (%)	Precision (%)	Specificity (%)	F_1	Balanced Accuracy (%)
Travel to log	78.42	51.68	91.63	0.62	85.02
Set choker	83.59	57.82	91.36	0.68	87.48
Travel away	83.06	41.42	94.73	0.55	88.90
Clear	76.95	97.08	93.72	0.86	85.33

https://doi.org/10.1371/journal.pone.0250624.t004

Table 5. Confusion matrix for the best choker setter model (created with a 3-s window and 90% overlap).

		Actual			
		Set choker	Travel away	Travel to log	Clear
Predicted	Set choker	18655	455	1898	11254
	Travel away	888	6409	554	7622
	Travel to log	1743	365	14455	11407
	Clear	1031	487	1526	101095

https://doi.org/10.1371/journal.pone.0250624.t005

combination of this variability with the large proportion of time workers spent in the clear likely reduced prediction accuracy and could account for the generally low precision and F_1 values. Since this form of analysis is very new in natural resources work, there is relatively little prior research on how delay, misclassified activities, or low precision in general may impact real-time summaries of work, or worker health and safety considerations, when interpreted in practice. Rather than a shortcoming of our study, we feel this is an important consideration for subsequent research and for development of wearable-based digital health and safety analytics for operational forestry. Additionally, conventional time study analysis distinguishes between quantifying individual productive cycle elements and quantifying overall productive and delay components of occupational work. Future modeling studies may benefit from distinguishing sampling efforts to quantify these separately. This may influence the way in which low precision and high false positive rates affect metrics such as productivity that may be calculated from model predictions in real-time. Thus, future work should consider how *clear* is defined and whether delay times are included in sampling and developed models in an effort to improve precision. Future work should also address issues of low precision through improved data collection, processing, and modeling (e.g., utilizing video recording, faster sampling rates, different sensors, deep learning, additional features, etc.), all of which are discussed below.

The best model performance metrics for both workers were obtained using the highest level of window overlap (90%) and smaller window sizes. The increasing accuracy we obtained with higher levels of window overlap is consistent with previous work [45]. However, many wrist-based models developed with either no overlap [27, 29, 56] or 50% overlap [54, 55] have achieved high accuracies. In terms of window size, previous work has suggested that simpler, more repetitive activities may be accurately captured with shorter windows, while more complex, less repetitive activities may need longer windows [34, 55]. In general, we observed slightly decreasing performance metrics with increasing window size. While some previous studies have found improved accuracies with increasing window size [34, 56], others have observed that increasing window size did not result in significant model improvement [55] and have developed accurate wrist-based models using smaller window sizes [35, 54, 55].

Smartwatches, smartphones, and other wearables generally have limited resources in terms of battery power, memory, storage, and computational power, so the effects of window size, window overlap, and sampling rate must be considered when designing and implementing occupational activity recognition models [31, 42, 111, 112]. For instance, using longer

Table 6. Accuracy metrics for the best chaser model (created with a 1-s window and 90% overlap).

Activity	Sensitivity (%)	Precision (%)	Specificity (%)	F_1	Balanced Accuracy (%)
Travel to log	78.53	14.74	92.43	0.25	85.48
Unhook	82.75	20.66	90.34	0.33	86.54
Travel away	81.93	16.75	89.76	0.28	85.84
Clear	71.95	99.16	91.99	0.83	81.97

https://doi.org/10.1371/journal.pone.0250624.t006

Table 7. Confusion matrix for the best chaser model (created with a 1-s window and 90% overlap).

		Actual			
		Clear	Travel away	Travel to log	Unhook
Predicted	Clear	366814	972	1022	1099
	Travel away	53458	11028	372	978
	Travel to log	39715	409	7061	714
	Unhook	49828	1052	536	13385

https://doi.org/10.1371/journal.pone.0250624.t007

windows [45, 111], higher levels of window overlap [45], and faster sampling rates [31, 42] requires more computational resources during implementation. In our study, the difference between the 90% and 75% overlap levels was noticeable for many of the model accuracy metrics. While lower levels of overlap may be preferable for real-time implementation on resource-limited devices such as smartwatches, the improved accuracies observed with the higher overlap levels seem to justify the additional computational power needed. On the other hand, our results suggest that the highest accuracies for both the choker setter and chaser models can be obtained with smaller window sizes, which indicates real-time implementation may be achieved with relatively minimal computational complexity and fast detection rates. Finally, future work should evaluate the effects of different sampling rates. While previous work has shown that higher frequencies can lead to improved accuracies [42, 111], others have achieved high performance using lower sampling rates such as 2-Hz [111] and 10-Hz [30]. The RawLogger application used to record the watch sensor data in this study did not allow us to adjust the sampling frequency, and Garmin Connect allows a maximum rate of 25-Hz for the accelerometer. However, different devices could be used in future work to assess whether higher frequencies may improve performance and whether lower sampling frequency may achieve similar performance, as this would have the benefit of reducing power consumption on the device.

Other considerations related to performance and resource consumption on wearable devices are the types of features used for classification and the tradeoffs between classical machine learning and deep learning methods. We chose to use time domain features primarily because they are less computationally complex and consume less energy during implementation [42, 113]. Additionally, it has been shown that the use of frequency domain features should be paired with faster sampling rates in order to achieve high model performance [42]. However, the use of frequency domain features should be investigated as a potential method of improving model performance [114]. Similarly, we chose to use a traditional machine learning approach because deep learning methods are generally more computationally expensive [47]. As devices become more powerful, deep learning offers opportunities to overcome some of the limitations of classical machine learning [48–50].

Activity recognition models based on wearables such as smartphones and smartwatches have practical applications for quantifying productivity and reporting work analytics to support digital health and safety. In terms of rigging crew productivity, models may inform the real-time reporting of productive work elements such as mean time per turn, number of chokers set or disconnected per turn, and related analytics. Models developed using conventional time studies in forestry have been limited to very specific operations and time periods under consideration, or deployed broadly in generalized predictive models (e.g., Bell et al. [115]). Utilizing our activity recognition models and subsequent, improved models using similar approaches makes it possible to quantify work day-in, day-out over the course of the work week and indefinitely in the future when incorporated into predictive apps on smartwatches. This is a major advancement and opens new possibilities for analysis of big data accumulated over time. For example, when summarized at the individual and group level, this information

could be used to improve productivity, reduce costs, and enhance work quality and worker safety by allowing workers to adapt treatment methods in near real-time. Furthermore, in broader meta-analysis, use of wearables to quantify productivity with greater temporal resolution than traditional time study techniques may provide opportunities for improving estimates of work productivity and treatment costs across various stand and site conditions and spatial scales. To better quantify worker health and safety, activity recognition model predictions may be paired simultaneously with fall detection algorithms to inform smart alerts indicating lack of movement associated with potential incidents. These alerts may then be sent to coworkers at the jobsite using emerging technologies designed to facilitate communication and data-sharing to improve safety in remote work areas. Model predictions may also be combined with other health-related data, such as heart rate, heat stress, and sleep metrics, to develop alerts that would be triggered when demanding work activities occur in conjunction with high physical exertion levels, increased heat stress, excess workload, or a combination of these factors. Finally, pairing activity recognition model predictions with real-time location information may be used to inform safety alerts related to proximity to coworkers, jobsite hazards such as snags or falling trees, and heavy equipment. This information could also be used in post-hoc analyses to better characterize the activities and other location- and health-related factors preceding accidents and near miss events.

One limitation to our study was reduced visibility when observing workers. Choker setters commonly work on steep hillsides, often among shrubs and other vegetation. At times, this made it difficult to visually observe start and stop times for component work cycle elements with precision. Chasers working at the landing were occasionally difficult to see when the yarder's movement shielded the chaser from view. A related limitation was a slight difference between the internal clocks of the smartwatches and the phones used to record the start and stop times for each activity. While these differences were small, there was no practical way to correct them in the field as they occasionally shift within a 24-hour period. This introduced additional error when assigning labels to the watch sensor data and is an important consideration in future remote occupational digital health applications, particularly for mobile and wearable devices that sync National Institute of Standards and Technology (NIST) time via internet connectivity.

Another limitation to the study was the short duration of many of the work activity cycles, primarily for the chaser. The chaser often runs in to and away from the yarded logs (both of which may last only a few seconds), and unhooking can be very quick (i.e., 3–5 s). This makes it difficult to visually detect and record accurate start and stop times for the rapid cycle elements that result. While this may explain why the shorter window lengths generally resulted in higher accuracies, it also made the activities more difficult to capture and model. Development of subsequent wearable- and mobile-based activity recognition for occupational safety may benefit from use of video concurrently with smartwatch sensor data collection rather than direct visual observation, in order to improve work element detection and support model training. For example, use of body camera videography coupled with post-hoc analysis to quantify work elements for low visibility tasks in forestry may improve model development.

A few areas that should be considered in future research include evaluating additional sensors, utilizing devices on different body parts, and incorporating more mechanistic approaches to modeling. The Garmin watches used in this study were only capable of recording raw accelerometer data. However, the incorporation of gyroscope sensor measurements into model development in future studies could potentially help to strengthen predictive power. While previous research has indicated that watch accelerometers perform better than watch gyroscopes for activity recognition [29], others have found that using both accelerometers and gyroscopes increases accuracy for smartphone-based models [52] and watch-based models

[32]. Additionally, previous studies have shown that hip-mounted accelerometers may recognize activities like running better than wrist-mounted accelerometers [27] and that using sensors in multiple positions, such as the wrist and pocket, can improve model performance [34]. Thus, future work should evaluate how placing devices on different body parts affects model predictions. For instance, it is possible that using sensors on the torso or feet may more accurately recognize the movements of choker setters and chasers as they travel to and from logs, while wrist-based sensors may be better suited to detecting activities that involve hand motions, like setting and disconnecting chokers. Another consideration when evaluating device locations is that previous research has shown that participants may prefer wrist-worn devices compared to hip-worn devices [116], which has implications for designing a relatively unobtrusive system for real-world adoption. Lastly, use of the random forest machine learning algorithm to model the occupational activities of rigging crew workers is highly empirical and doesn't necessarily help to foster understanding of the underlying processes affecting work productivity and safety. In future studies, a more mechanistic approach to modeling work movements may better lend itself to identifying causal relationships associated with safety incidents and possible interventions.

Because logging operations are highly variable and our sampling was observational, the amount of data collected from each participant varied. We chose to randomly separate the entire dataset into 2/3 training and 1/3 testing since utilizing either leave-one-subject-out or k-fold cross-validation implemented at the participant level would have meant creating models with varying quantities of data in each iteration. Thus, because our validation methods may randomly include data from participants in both training and testing, the quality of predictive models presented may be overly optimistic. Future research developing similar models may benefit from a different approach that avoids cross-over of participant data in training and testing subsets. Additionally, traditional time and motion and actigraphy analysis in forestry work has generally been based on relatively small studies. Because use of IMU sensors to quantify work in real-time is relatively new, it is unknown whether the mean and variability of data in our study would be fully representative of the broad range of field sites, forest stand conditions, equipment, and weather impacts that affect worker movements in the profession overall. While we believe the quantity of data from each participant used to train and test the models was sufficient, future studies may benefit from collecting a more balanced sample of data from a wider variety of participants. Our goals in this study were to evaluate, at a broad level, the potential for wearable devices to model real-time occupational rigging crew work activities and to provide an example of the methods, modeling approaches, and sampling considerations that are important for developing libraries of generalized forestry work activity recognition models. Prior to use in occupational settings, predictions from real-time models developed, regardless of the statistical validation methodology used in model fitting and analysis, should be further evaluated using data collected independently as part of different field operations reflecting variability in site conditions, weather, workers, and other factors.

Ultimately, future work should include coding the best models developed in this study into a smartwatch application to support real-time characterization of work activities and further validate model predictions in a variety of conditions. Pairing of fall detection with activity recognition model predictions may help to inform development of improved smart alerts to coworkers notifying them of potential jobsite incidents, particularly when paired with real-time GNSS mapping in remote forestry work environments. To advance digital health and safety more broadly, the data resulting from our predictive models, as well as from models developed subsequently for other common forestry work tasks, may be used to quantify day-to-day occupational forestry job functions in high resolution. The resulting work effort data provide a fundamental mechanism through which it may be possible to better quantify factors

associated with incident occurrence across forest stand, site, weather, air quality, and other work conditions, particularly when paired with readily available, wearable-based personal health metrics such as sleep activity, heart rate, and heat stress.

Acknowledgments

The authors would like to thank participating state and regional industrial timberland owners for permitting us to conduct operational sampling on their timber sales. The authors would also like to thank 14 anonymous loggers who supported our research by wearing smartwatches and agreeing to allow us to observe their daily operations.

Author Contributions

Conceptualization: Eloise G. Zimbelman, Robert F. Keefe.

Data curation: Eloise G. Zimbelman.

Formal analysis: Eloise G. Zimbelman, Robert F. Keefe.

Funding acquisition: Eloise G. Zimbelman, Robert F. Keefe.

Investigation: Eloise G. Zimbelman, Robert F. Keefe.

Methodology: Eloise G. Zimbelman, Robert F. Keefe.

Project administration: Eloise G. Zimbelman, Robert F. Keefe.

Resources: Eloise G. Zimbelman, Robert F. Keefe.

Software: Eloise G. Zimbelman.

Supervision: Eloise G. Zimbelman, Robert F. Keefe.

Validation: Eloise G. Zimbelman.

Visualization: Eloise G. Zimbelman.

Writing – original draft: Eloise G. Zimbelman.

Writing – review & editing: Eloise G. Zimbelman, Robert F. Keefe.

References

1. Hartley DS, Han H-S. Effects of alternative silvicultural treatments on cable harvesting productivity and cost in western Washington. West J Appl For. 2007; 22: 204–212. https://doi.org/10.1093/wjaf/22.3.204
2. Huyler NK, LeDoux CB. Cycle-time equation for the Koller K300 cable yarder operating on steep slopes in the Northeast. Radnor, PA, USA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station; 1997 Mar. Report No.: Res. Pap. NE-705. https://doi.org/10.2737/NE-RP-705
3. Kirk PM, Sullman MJM. Heart rate strain in cable hauler choker setters in New Zealand logging operations. Appl Ergon. 2001; 32: 389–398. https://doi.org/10.1016/s0003-6870(01)00003-5 PMID: 11461040
4. Shaffer RM, Milburn JS. Injuries on feller-buncher/grapple skidder logging operations in the Southeastern United States. For Prod J. 1999; 49: 24–26.
5. Bell JL. Changes in logging injury rates associated with use of feller-bunchers in West Virginia. J Safety Res. 2002; 33: 463–471. https://doi.org/10.1016/s0022-4375(02)00048-8 PMID: 12429103
6. Lefort AJ Jr, de Hoop CF, Pine JC, Marx BD. Characteristics of injuries in the logging industry of Louisiana, USA: 1986 to 1998. Int J For Eng. 2003; 14: 75–89. https://doi.org/10.1080/14942119.2003.10702480
7. Scott DF. A study of logger fatalities from 1992–2000. Inj Prev. 2004; 10: 239–243. https://doi.org/10.1136/ip.2003.004663 PMID: 15314053

8. Lagerstrom E, Magzamen S, Rosecrance J. A mixed-methods analysis of logging injuries in Montana and Idaho. Am J Ind Med. 2017; 60: 1077–1087. https://doi.org/10.1002/ajim.22759 PMID: 28833349

9. U.S. Bureau of Labor Statistics. Fatal occupational injuries by occupation and event or exposure, all United States, 2017. 2018. Available from: https://www.bls.gov/iif/oshwc/cfoi/cftb0317.htm

10. U.S. Bureau of Labor Statistics. Fatal occupational injuries by occupation and event or exposure, all United States, 2018. 2019. Available from: https://www.bls.gov/iif/oshwc/cfoi/cftb0326.htm

11. Tsioras PA, Rottensteiner C, Stampfer K. Analysis of accidents during cable yarding operations in Austria 1998–2008. Croat J For Eng. 2011; 32: 549–560.

12. Bordas RM, Davis GA, Hopkins BL, Thomas RE, Rummer RB. Documentation of hazards and safety perceptions for mechanized logging operations in East Central Alabama. J Agric Saf Health. 2001; 7: 113–123. https://doi.org/10.13031/2013.2673 PMID: 11465386

13. Wempe AM, Keefe RF. Characterizing rigging crew proximity to hazards on cable logging operations using GNSS-RF: effect of GNSS positioning error on worker safety status. Forests. 2017; 8: 357. https://doi.org/10.3390/f8100357 PMID: 30976381

14. Zimbelman EG, Keefe RF, Strand EK, Kolden CA, Wempe AM. Hazards in motion: Development of mobile geofences for use in logging safety. Sensors. 2017; 17: 822. https://doi.org/10.3390/s17040822 PMID: 28394303

15. Newman SM, Keefe RF, Brooks RH, Ahonen EQ, Wempe AM. Human factors affecting logging injury incidents in Idaho and the potential for real-time location-sharing technology to improve safety. Safety. 2018; 4: 43. https://doi.org/10.3390/safety4040043 PMID: 30515383

16. Zimbelman EG, Keefe RF. Real-time positioning in logging: Effects of forest stand characteristics, topography, and line-of-sight obstructions on GNSS-RF transponder accuracy and radio signal propagation. PLoS ONE. 2018; 13: e0191017. https://doi.org/10.1371/journal.pone.0191017 PMID: 29324794

17. Wempe AM, Keefe RF, Newman SM, Paveglio TB. Intent to adopt location sharing for logging safety applications. Safety. 2019; 5: 7. https://doi.org/10.3390/safety5010007

18. Bao L, Intille SS. Activity recognition from user-annotated acceleration data. In: Ferscha A, Mattern F, editors. Proceedings of the Second International Conference on Pervasive Computing; 2004 Apr 21–23; Linz/Vienna, Austria. Springer, Berlin, Heidelberg; 2004. pp. 1–17. https://doi.org/10.1007/978-3-540-24646-6_1

19. Mathie MJ, Celler BG, Lovell NH, Coster ACF. Classification of basic daily movements using a triaxial accelerometer. Med Biol Eng Comput. 2004; 42: 679–687. https://doi.org/10.1007/BF02347551 PMID: 15503970

20. Wu W, Dasgupta S, Ramirez EE, Peterson C, Norman GJ. Classification accuracies of physical activities using smartphone motion sensors. J Med Internet Res. 2012; 14: e130. https://doi.org/10.2196/jmir.2208 PMID: 23041431

21. Micucci D, Mobilio M, Napoletano P. UniMiB SHAR: a dataset for human activity recognition using acceleration data from smartphones. Appl Sci. 2017; 7: 1101. https://doi.org/10.3390/app7101101

22. Ermes M, Pärkkä J, Mäntyjärvi J, Korhonen I. Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. IEEE Trans Inf Technol Biomed. 2008; 12: 20–26. https://doi.org/10.1109/TITB.2007.899496 PMID: 18270033

23. Mitchell E, Monaghan D, O'Connor NE. Classification of sporting activities using smartphone accelerometers. Sensors. 2013; 13: 5317–5337. https://doi.org/10.3390/s130405317 PMID: 23604031

24. Stöggl T, Holst A, Jonasson A, Andersson E, Wunsch T, Norström C, et al. Automatic classification of the sub-techniques (gears) used in cross-country ski skating employing a mobile phone. Sensors. 2014; 14: 20589–20601. https://doi.org/10.3390/s141120589 PMID: 25365459

25. Lane ND, Miluzzo E, Lu H, Peebles D, Choudhury T, Campbell AT. A survey of mobile phone sensing. IEEE Commun Mag. 2010; 48: 140–150. https://doi.org/10.1109/MCOM.2010.5560598

26. Kwapisz JR, Weiss GM, Moore SA. Activity recognition using cell phone accelerometers. ACM SIGKDD Explor Newsl. 2011; 12: 74–82. https://doi.org/10.1145/1964897.1964918

27. Trost SG, Zheng Y, Wong W-K. Machine learning for activity recognition: hip versus wrist data. Physiol Meas. 2014; 35: 2183–2189. https://doi.org/10.1088/0967-3334/35/11/2183 PMID: 25340887

28. del Rosario MB, Redmond SJ, Lovell NH. Tracking the evolution of smartphone sensing for monitoring human movement. Sensors. 2015; 15: 18901–18933. https://doi.org/10.3390/s150818901 PMID: 26263998

29. Weiss GM, Timko JL, Gallagher CM, Yoneda K, Schreiber AJ. Smartwatch-based activity recognition: A machine learning approach. Proceedings of the 2016 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI); 2016 Feb 24–27; Las Vegas, NV, USA. IEEE; 2016. pp. 426–429. https://doi.org/10.1109/BHI.2016.7455925

30. Keefe RF, Zimbelman EG, Wempe AM. Use of smartphone sensors to quantify the productive cycle elements of hand fallers on industrial cable logging operations. Int J For Eng. 2019; 30: 132–143. https://doi.org/10.1080/14942119.2019.1572489

31. Shoaib M, Bosch S, Incel OD, Scholten H, Havinga PJM. A survey of online activity recognition using mobile phones. Sensors. 2015; 15: 2059–2085. https://doi.org/10.3390/s150102059 PMID: 25608213

32. Mortazavi BJ, Pourhomayoun M, Alsheikh G, Alshurafa N, Lee SI, Sarrafzadeh M. Determining the single best axis for exercise repetition recognition and counting on smartwatches. Proceedings of the 2014 11th International Conference on Wearable and Implantable Body Sensor Networks; 2014 Jun 16–19; Zurich, Switzerland. IEEE; 2014. pp. 33–38. https://doi.org/10.1109/BSN.2014.21

33. Moschetti A, Fiorini L, Esposito D, Dario P, Cavallo F. Recognition of daily gestures with wearable inertial rings and bracelets. Sensors. 2016; 16: 1341. https://doi.org/10.3390/s16081341 PMID: 27556473

34. Shoaib M, Bosch S, Incel OD, Scholten H, Havinga PJM. Complex human activity recognition using smartphone and wrist-worn motion sensors. Sensors. 2016; 16: 426. https://doi.org/10.3390/s16040426 PMID: 27023543

35. Cvetković B, Szeklicki R, Janko V, Lutomski P, Luštrek M. Real-time activity monitoring with a wristband and a smartphone. Inf Fusion. 2018; 43: 77–93. https://doi.org/10.1016/j.inffus.2017.05.004

36. Lara OD, Labrador MA. A survey on human activity recognition using wearable sensors. IEEE Commun Surv Tutor. 2013; 15: 1192–1209. https://doi.org/10.1109/SURV.2012.110112.00192

37. Saha J, Chowdhury C, Roy Chowdhury I, Biswas S, Aslam N. An ensemble of condition based classifiers for device independent detailed human activity recognition using smartphones. Information. 2018; 9: 94. https://doi.org/10.3390/info9040094

38. Chen Y, Shen C. Performance analysis of smartphone-sensor behavior for human activity recognition. IEEE Access. 2017; 5: 3095–3110. https://doi.org/10.1109/ACCESS.2017.2676168

39. Karantonis DM, Narayanan MR, Mathie M, Lovell NH, Celler BG. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Trans Inf Technol Biomed. 2006; 10: 156–167. https://doi.org/10.1109/titb.2005.856864 PMID: 16445260

40. Bayat A, Pomplun M, Tran DA. A study on human activity recognition using accelerometer data from smartphones. Procedia Comput Sci. 2014; 34: 450–457. https://doi.org/10.1016/j.procs.2014.07.009

41. Khan AM, Tufail A, Khattak AM, Laine TH. Activity recognition on smartphones via sensor-fusion and KDA-based SVMs. Int J Distrib Sens Netw. 2014; 10: 503291. https://doi.org/10.1155/2014/503291

42. Khan AM, Siddiqi MH, Lee S-W. Exploratory data analysis of acceleration signals to select light-weight and accurate features for real-time activity recognition on smartphones. Sensors. 2013; 13: 13099–13122. https://doi.org/10.3390/s131013099 PMID: 24084108

43. Attal F, Mohammed S, Dedabrishvili M, Chamroukhi F, Oukhellou L, Amirat Y. Physical human activity recognition using wearable sensors. Sensors. 2015; 15: 31314–31338. https://doi.org/10.3390/s151229858 PMID: 26690450

44. Davila JC, Cretu A-M, Zaremba M. Wearable sensor data classification for human activity recognition based on an iterative learning framework. Sensors. 2017; 17: 1287. https://doi.org/10.3390/s17061287 PMID: 28590422

45. Janidarmian M, Roshan Fekr A, Radecka K, Zilic Z. A comprehensive analysis on wearable acceleration sensors in human activity recognition. Sensors. 2017; 17: 529. https://doi.org/10.3390/s17030529 PMID: 28272362

46. Mehrang S, Pietilä J, Korhonen I. An activity recognition framework deploying the random forest classifier and a single optical heart rate monitoring and triaxial accelerometer wrist-band. Sensors. 2018; 18: 613. https://doi.org/10.3390/s18020613 PMID: 29470385

47. Bhattacharya S, Lane ND. From smart to deep: Robust activity recognition on smartwatches using deep learning. Proceedings of the 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops); 2016 Mar 14–18; Sydney, NSW, Australia. IEEE; 2016. pp. 1–6. https://doi.org/10.1109/PERCOMW.2016.7457169

48. Ronao CA, Cho S-B. Human activity recognition with smartphone sensors using deep learning neural networks. Expert Syst Appl. 2016; 59: 235–244. https://doi.org/10.1016/j.eswa.2016.04.032

49. Nweke HF, Teh YW, Al-garadi MA, Alo UR. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. Expert Syst Appl. 2018; 105: 233–261. https://doi.org/10.1016/j.eswa.2018.03.056

50. Wang J, Chen Y, Hao S, Peng X, Hu L. Deep learning for sensor-based activity recognition: A survey. Pattern Recognit Lett. 2019; 119: 3–11. https://doi.org/10.1016/j.patrec.2018.02.010

51. Banos O, Galvez J-M, Damas M, Pomares H, Rojas I. Window size impact in human activity recognition. Sensors. 2014; 14: 6474–6499. https://doi.org/10.3390/s140406474 PMID: 24721766

52. Dernbach S, Das B, Krishnan NC, Thomas BL, Cook DJ. Simple and complex activity recognition through smart phones. Proceedings of the 2012 Eighth International Conference on Intelligent Environments; 2012 Jun 26–29; Guanajuato, Mexico. IEEE; 2012. pp. 214–221. https://doi.org/10.1109/IE.2012.39

53. Siirtola P, Röning J. Ready-to-use activity recognition for smartphones. Proceedings of the 2013 IEEE Symposium on Computational Intelligence and Data Mining (CIDM); 2013 Apr 16–19; Singapore, Singapore. IEEE; 2013. pp. 59–64. https://doi.org/10.1109/CIDM.2013.6597218

54. Shoaib M, Bosch S, Incel OD, Scholten H, Havinga PJM. Fusion of smartphone motion sensors for physical activity recognition. Sensors. 2014; 14: 10146–10176. https://doi.org/10.3390/s140610146 PMID: 24919015

55. Shoaib M, osch S, Scholten H, Havinga PJM, Incel OD. Towards detection of bad habits by fusing smartphone and smartwatch sensors. Proceedings of the 2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops); 2015 Mar 23–27; St. Louis, MO, USA. IEEE; 2015. pp. 591–596. https://doi.org/10.1109/PERCOMW.2015.7134104

56. Mannini A, Intille SS, Rosenberger M, Sabatini AM, Haskell W. Activity recognition using a single accelerometer placed at the wrist or ankle. Med Sci Sports Exerc. 2013; 45: 2193–2203. https://doi.org/10.1249/MSS.0b013e31829736d6 PMID: 23604069

57. Miluzzo E, Lane ND, Fodor K, Peterson R, Lu H, Musolesi M, et al. Sensing meets mobile social networks: the design, implementation and evaluation of the CenceMe application. Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems (SenSys '08); 2008 Nov 5–7; Raleigh, NC, USA. ACM; 2008. pp. 337–350. https://doi.org/10.1145/1460412.1460445

58. Reddy S, Mun M, Burke J, Estrin D, Hansen M, Srivastava M. Using mobile phones to determine transportation modes. ACM Trans Sens Netw. 2010; 6: 13. https://doi.org/10.1145/1689239.1689243

59. Ward JA, Lukowicz P, Tröster G, Starner TE. Activity recognition of assembly tasks using body-worn microphones and accelerometers. IEEE Trans Pattern Anal Mach Intell. 2006; 28: 1553–1567. https://doi.org/10.1109/TPAMI.2006.197 PMID: 16986539

60. Stiefmeier T, Roggen D, Ogris G, Lukowicz P, Tröster G. Wearable activity tracking in car manufacturing. IEEE Pervasive Comput. 2008; 7: 42–50. https://doi.org/10.1109/MPRV.2008.40

61. Joshua L, Varghese K. Accelerometer-based activity recognition in construction. J Comput Civ Eng. 2011; 25: 370–379. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000097

62. Akhavian R, Behzadan AH. Smartphone-based construction workers' activity recognition and classification. Autom Constr. 2016; 71: 198–209. https://doi.org/10.1016/j.autcon.2016.08.015

63. Valero E, Sivanathan A, Bosché F, Abdel-Wahab M. Musculoskeletal disorders in construction: A review and a novel system for activity tracking with body area network. Appl Ergon. 2016; 54: 120–130. https://doi.org/10.1016/j.apergo.2015.11.020 PMID: 26851471

64. Sánchez D, Tentori M, Favela J. Activity recognition for the smart hospital. IEEE Intell Syst. 2008; 23: 50–57. https://doi.org/10.1109/MIS.2008.18

65. Zhao W, Lun R, Gordon C, Fofana A-BM, Espy DD, Reinthal MA, et al. A human-centered activity tracking system: toward a healthier workplace. IEEE Trans Hum-Mach Syst. 2017; 47: 343–355. https://doi.org/10.1109/THMS.2016.2611825

66. Pierzchała M, Kvaal K, Stampfer K, Talbot B. Automatic recognition of work phases in cable yarding supported by sensor fusion. Int J For Eng. 2018; 29: 12–20. https://doi.org/10.1080/14942119.2017.1373502

67. Barnes RM. Motion and time study. 4th ed. New York: Wiley; 1958.

68. Olsen ED, Kellogg LD. Comparison of time-study techniques for evaluating logging production. Trans ASAE. 1983; 26: 1665–1668. https://doi.org/10.13031/2013.33821

69. Lortz D, Kluender R, McCoy W. Manual felling time and productivity in southern pine forests. For Prod J. 1997; 47: 59–63.

70. Wang J, McNeel J, Baumgras J. A computer-based time study system for timber harvesting operations. For Prod J. 2003; 53: 47–53.

71. Adebayo AB, Han H-S, Johnson L. Productivity and cost of cut-to-length and whole-tree harvesting in a mixed-conifer stand. For Prod J. 2007; 57: 59–69.

72. Spinelli R, Visser R. Analyzing and estimating delays in harvester operations. Int J For Eng. 2008; 19: 36–41. https://doi.org/10.1080/14942119.2008.10702558

73. Magagnotti N, Kanzian C, Schulmeyer F, Spinelli R. A new guide for work studies in forestry. Int J For Eng. 2013; 24: 249–253. https://doi.org/10.1080/14942119.2013.856613

74. Wang J, Long C, McNeel J, Baumgras J. Productivity and cost of manual felling and cable skidding in central Appalachian hardwood forests. For Prod J. 2004; 54: 45–51.

75. Spinelli R, Nati C, Magagnotti N. Using modified foragers to harvest short-rotation poplar plantations. Biomass Bioenergy. 2009; 33: 817–821. https://doi.org/10.1016/j.biombioe.2009.01.001

76. Rossit DA, Olivera A, Viana Céspedes V, Broz D. A Big Data approach to forestry harvesting productivity. Comput Electron Agric. 2019; 161: 29–52. https://doi.org/10.1016/j.compag.2019.02.029

77. Proto AR, Sperandio G, Costa C, Maesano M, Antonucci F, Macrì G, et al. A three-step neural network artificial intelligence modeling approach for time, productivity and costs prediction: a case study in Italian forestry. Croat J For Eng. 2020; 41: 35–47. https://doi.org/10.5552/crojfe.2020.611

78. Miyata ES. Determining fixed and operating costs of logging equipment. St. Paul, MN, USA: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station; 1980. Report No.: General Technical Report NC-55. Available from: https://www.nrs.fs.fed.us/pubs/gtr/gtr_nc055.pdf

79. Brinker RW, Kinard J, Rummer R, Lanford B. Machine rates for selected forest harvesting machines. Auburn, AL, USA: Alabama Agricultural Experiment Station; 2002 Sep. Report No.: Circular 296 (revised). Available from: https://www.fs.usda.gov/treesearch/pubs/33467

80. Bolding MC, Lanford BL. Wildfire fuel harvesting and resultant biomass utilization using a cut-to-length/small chipper system. For Prod J. 2005; 55: 181–189.

81. Strandgard M, Mitchell R. Automated time study of forwarders using GPS and a vibration sensor. Croat J For Eng. 2015; 36: 175–184.

82. Olivera A, Visser R, Acuna M, Morgenroth J. Automatic GNSS-enabled harvester data collection as a tool to evaluate factors affecting harvester productivity in a Eucalyptus spp. harvesting operation in Uruguay. Int J For Eng. 2016; 27: 15–28. https://doi.org/10.1080/14942119.2015.1099775

83. Becker RM, Keefe RF, Anderson NM. Use of real-time GNSS-RF data to characterize the swing movements of forestry equipment. Forests. 2017; 8: 44. https://doi.org/10.3390/f8020044

84. Gallo R, Grigolato S, Cavalli R, Mazzetto F. GNSS-based operational monitoring devices for forest logging operation chains. J Agric Eng. 2013; 44: 140–144. https://doi.org/10.4081/jae.2013.269

85. McDonald TP, Fulton JP, Darr MJ, Gallagher TV. Evaluation of a system to spatially monitor hand planting of pine seedlings. Comput Electron Agric. 2008; 64: 173–182. https://doi.org/10.1016/j.compag.2008.04.011

86. Borz SA, Talagai N, Cheța M, Gavilanes Montoya AV, Castillo Vizuete DD. Automating data collection in motor-manual time and motion studies implemented in a willow short rotation coppice. BioResources. 2018; 13: 3236–3249.

87. Talagai N, Cheța M, Gavilanes Montoya AV, Castillo Vizuete DD, Borz SA. Predicting time consumption of chipping tasks in a willow short rotation coppice from GPS and acceleration data. Proceedings of the 8th Edition of the Biennial International Symposium: "Forest and Sustainable Development"; 2018 Oct 25–27; Bra□ov, Romania. Transilvania University Press; 2018. pp. 1–12.

88. Cadei A, Mologni O, Proto AR, D'Anna G, Grigolato S. Using high-frequency accelerometer to detect machine tilt. Proceedings of the 19th International Scientific Conference: Engineering for Rural Development; 2020 May 20–22; Jelgava, Latvia. Latvia University of Life Sciences and Technologies; 2020. pp. 1865–1870. https://doi.org/10.22616/ERDev2020.19.TF512

89. Bowen J, Hinze A, Cunningham SJ, Parker R. Evaluating low-cost activity trackers for use in large-scale data gathering of forestry workers. Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction (OzCHI '15); 2015 Dec 7–10; Parkville, Australia. ACM; 2015. pp. 474–482. https://doi.org/10.1145/2838739.2838741

90. Bowen J, Hinze A, Griffiths C. Investigating real-time monitoring of fatigue indicators of New Zealand forestry workers. Accid Anal Prev. 2019; 126: 122–141. https://doi.org/10.1016/j.aap.2017.12.010 PMID: 29287742

91. Cheța M, Marcu MV, Borz SA. Workload, exposure to noise, and risk of musculoskeletal disorders: a case study of motor-manual tree feeling and processing in poplar clear cuts. Forests. 2018; 9: 300. https://doi.org/10.3390/f9060300

92. Cheța M, Marcu MV, Iordache E, Borz SA. Testing the capability of low-cost tools and artificial intelligence techniques to automatically detect operations done by a small-sized manually driven bandsaw. Forests. 2020; 11: 739. https://doi.org/10.3390/f11070739

93. Cheța M, Marcu MV, Borz SA. Effect of training parameters on the ability of artificial neural networks to learn: a simulation on accelerometer data for task recognition in motor-manual felling and processing. Bull Transilv Univ Braşov. 2020; 13: 19–36. https://doi.org/10.31926/but.fwiafe.2020.13.62.1.2

94. m_c8bit. TimeStamp. m_c8bit; 2019. Version 0.4.0. Available from: https://play.google.com/store/apps/details?id=jp.m_c8bit.timestamp

95. Dufour C. RawLogger. 2019. Version 1.0.20190520a. Available from: http://apps.garmin.com/en-US/apps/bd37beed-35e1-4ffc-a9fc-9297723d6c72

96. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020. Available from: https://www.r-project.org/

97. Liaw A, Wiener M. Classification and regression by randomForest. R News. 2002; 2: 18–22.

98. Genuer R, Poggi J-M, Tuleau-Malot C. Variable selection using random forests. Pattern Recognit Lett. 2010; 31: 2225–2236. https://doi.org/10.1016/j.patrec.2010.03.014

99. Oshiro TM, Perez PS, Baranauskas JA. How many trees in a random forest? In: Perner P, editor. MLDM 2012: Machine Learning and Data Mining in Pattern Recognition; 2012 Jul 13–20; Berlin, Germany. Springer, Berlin, Heidelberg; 2012. pp. 154–168. https://doi.org/10.1007/978-3-642-31537-4_13

100. Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP. Random Forest: A classification and regression tool for compound classification and QSAR modeling. J Chem Inf Comput Sci. 2003; 43: 1947–1958. https://doi.org/10.1021/ci034160g PMID: 14632445

101. Kuhn M. caret: classification and regression training. 2020. R package version 6.0–85. Available from: https://CRAN.R-project.org/package=caret

102. Minnen D, Westeyn T, Starner T, Ward JA, Lukowicz P. Performance metrics and evaluation issues for continuous activity recognition. Proceedings of the Performance Metrics for Intelligent Systems Workshop; 2006 Aug 21–23; Gaithersburg, MD, USA. NIST; 2006. pp. 141–148.

103. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett. 2006; 27: 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

104. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982; 143: 29–36. https://doi.org/10.1148/radiology.143.1.7063747 PMID: 7063747

105. Bradley AP. The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern Recognit. 1997; 30: 1145–1159. https://doi.org/10.1016/S0031-3203(96)00142-2

106. Weiss GM, Provost F. Learning when training data are costly: the effect of class distribution on tree induction. J Artif Intell Res. 2003; 19: 315–354. https://doi.org/10.1613/jair.1199

107. Boughorbel S, Jarray F, El-Anbari M. Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. PLoS ONE. 2017; 12: e0177678. https://doi.org/10.1371/journal.pone.0177678 PMID: 28574989

108. Luque A, Carrasco A, Martín A, de las Heras A. The impact of class imbalance in classification performance metrics based on the binary confusion matrix. Pattern Recognit. 2019; 91: 216–231. https://doi.org/10.1016/j.patcog.2019.02.023

109. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011; 12: 77. https://doi.org/10.1186/1471-2105-12-77 PMID: 21414208

110. Hand DJ, Till RJ. A simple generalisation of the area under the ROC curve for multiple class classification problems. Mach Learn. 2001; 45: 171–186. https://doi.org/10.1023/A:1010920819831

111. Liang Y, Zhou X, Yu Z, Guo B. Energy-efficient motion related activity recognition on mobile devices for pervasive healthcare. Mob Netw Appl. 2014; 19: 303–317. https://doi.org/10.1007/s11036-013-0448-9

112. Rehman MH ur, Liew CS, Wah TY, Shuja J, Daghighi B. Mining personal data using smartphones and wearable devices: a survey. Sensors. 2015; 15: 4430–4469. https://doi.org/10.3390/s150204430 PMID: 25688592

113. Figo D, Diniz PC, Ferreira DR, Cardoso JMP. Preprocessing techniques for context recognition from accelerometer data. Pers Ubiquitous Comput. 2010; 14: 645–662. https://doi.org/10.1007/s00779-010-0293-9

114. Erdaş ÇB, Atasoy I, Açıcı K, Oğul H. Integrating features for accelerometer-based activity recognition. Procedia Comput Sci. 2016; 98: 522–527. https://doi.org/10.1016/j.procs.2016.09.070

115. Bell CK, Keefe RF, Fried JS. Validation of the OpCost logging cost model using contractor surveys. Int J For Eng. 2017; 28: 73–84. https://doi.org/10.1080/14942119.2017.1313488

116. Scott JJ, Rowlands AV, Cliff DP, Morgan PJ, Plotnikoff RC, Lubans DR. Comparability and feasibility of wrist- and hip-worn accelerometers in free-living adolescents. J Sci Med Sport. 2017; 20: 1101–1106. https://doi.org/10.1016/j.jsams.2017.04.017 PMID: 28501418